LA-UR-21-20971

Approved for public release; distribution is unlimited.

Title:	Subtask 3.1: Sequential Design of Experiments
Author(s):	Anderson-Cook, Christine Michaela Ng, Brenda
Intended for:	Report
Issued:	2021-02-03

CCSI²

Carbon Capture Simulation for Industry Impact

Subtask 3.1: Sequential Design of Experiments

Subtask Lead: Christine Anderson-Cook, LANL

Co-Lead: Brenda Ng, LLNL

January 2021

Subtask Team:

Pedro Sotorrio (LLNL), Abby Nachtsheim (LANL), Miranda Martin (LANL), Alex Dowling (Notre Dame), Jialu Wong (Notre Dame), Josh Morgan (NETL), Charles Tong (LLNL)

Why Design of Experiments in CCSI^2?

- Data in Carbon Capture applications are expensive
- Time at test facilities difficult to obtain and requires waiting

Want to make most of data that we are able to get

- Sequential Design of Experiments:
  - Allows us to be strategic about choosing what data are most beneficial
  - Tailor data collection to the specific goals of each experiment (or stage of experiment)
  - Leverage what we already know to take maximum advantage of new data
- SDoE module in FOQUS provides tools for experimenters to
  - Incorporate what is already known about a process
  - Quickly generate a designed experiment to match their objectives

EY20 Highlights

- New release of FOQUS SDoE module with new capabilities
  - **Space-Filling design**
    - Uniform Space-Filling (USF)
    - Non-Uniform Space-Filling (NUSF)
    - Input-Response Space-Filling (IRSF)
  - **Robust Optimality-Based design**
    - Leverages capabilities in UQ
    - Builds an empirical surrogate model
    - Construct design using G-, I-, D- or A-optimality

* New in EY20

Focus: good prediction

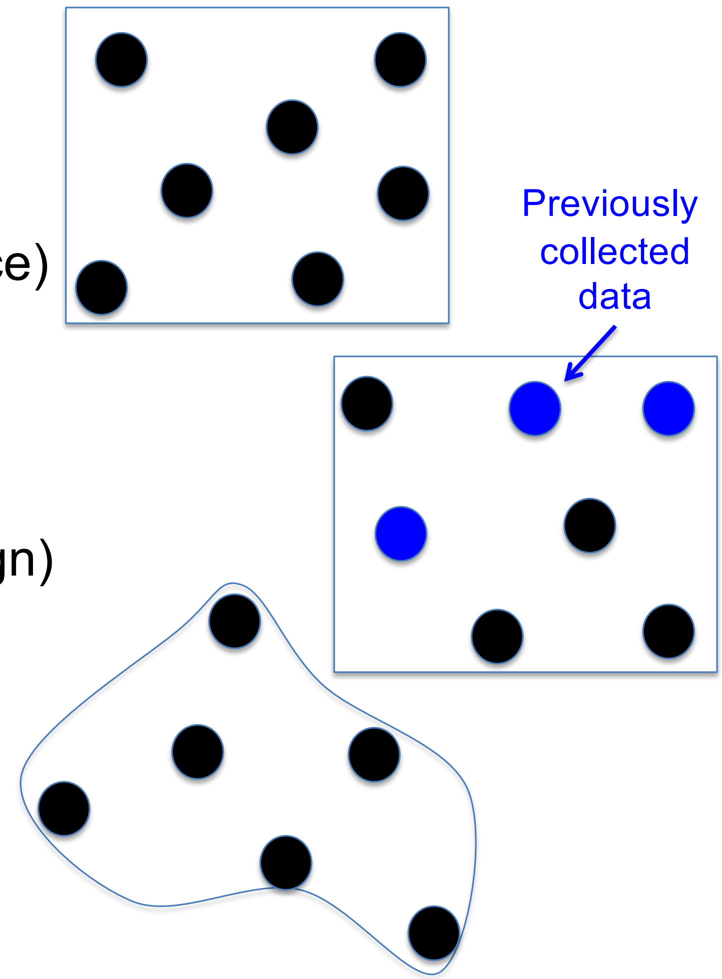
Focus: good estimation of model parameters

CCSI2 Carbon Capture Simulation for Industry Impact

Uniform Space-Filling Designs

- Inputs:
  - Candidate set (specifies dimension of input space)
  - Previous data (optional)
  - Minimax or Maximin
  - Size(s) of designs
  - Number of random starts (time to generate design)

- Outputs
  - Multiple designs with criteria values



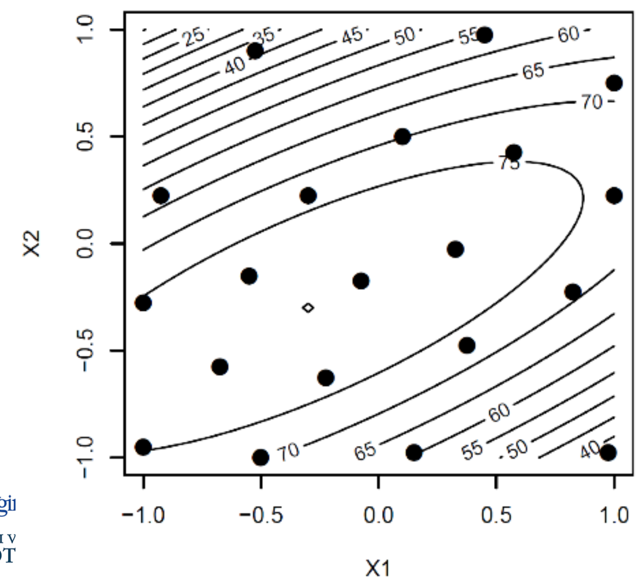
Non-Uniform Space-Filling Designs

- Inputs:
  - Candidate set (specifies dimension of input space)
  - Previous data (optional)
  - Size of design
  - MWR - Maximum Weight Ratio (degree of non-uniformity)
  - Direct or Ranked scaling of weights
  - Number of random starts (time to generate design)

Requires column for weights with value for each row

Flexible for different objectives

- Outputs
  - Multiple designs with criteria values

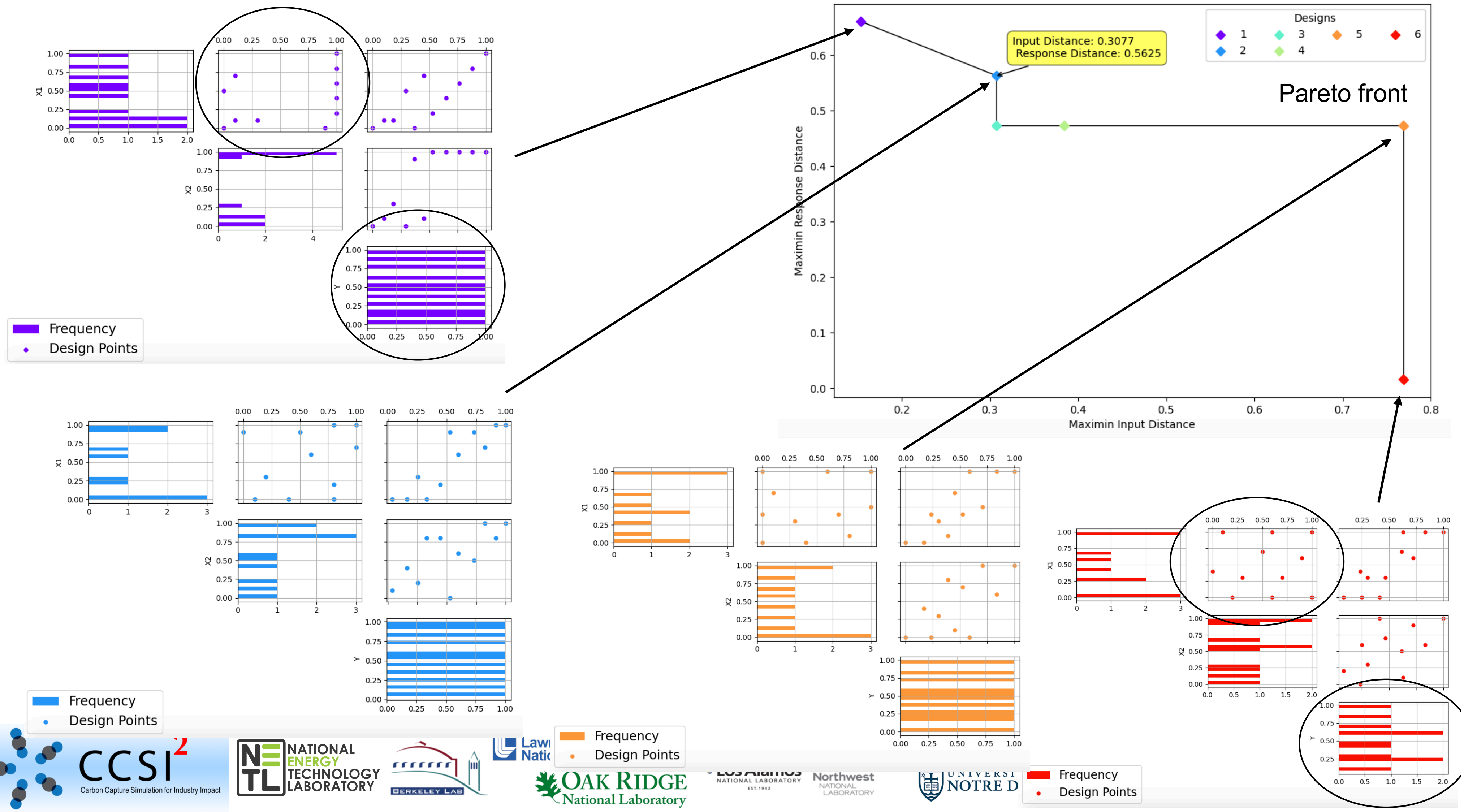


Input-Response Space-Filling Designs

- Inputs:
  - Candidate set (specifies dimension of input space)
  - Previous data (optional)
  - Minimax or Maximin
  - Size of design
  - Number of random starts (time to generate design)

Requires column for predicted response values

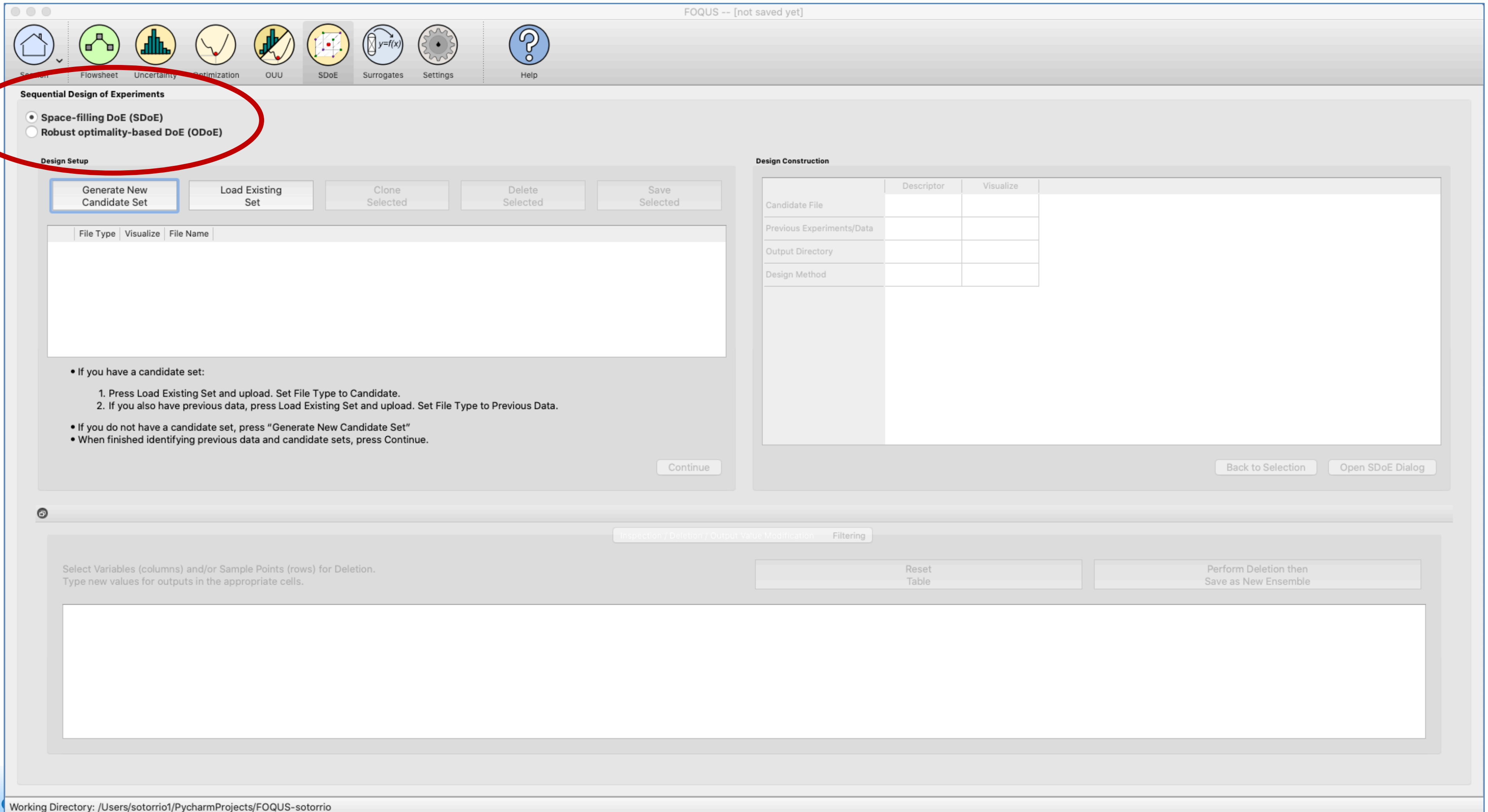
- Outputs
  - Pareto front of objectively best designs to balance spacing in input and response spaces
  - Details for each design – which runs and criteria values

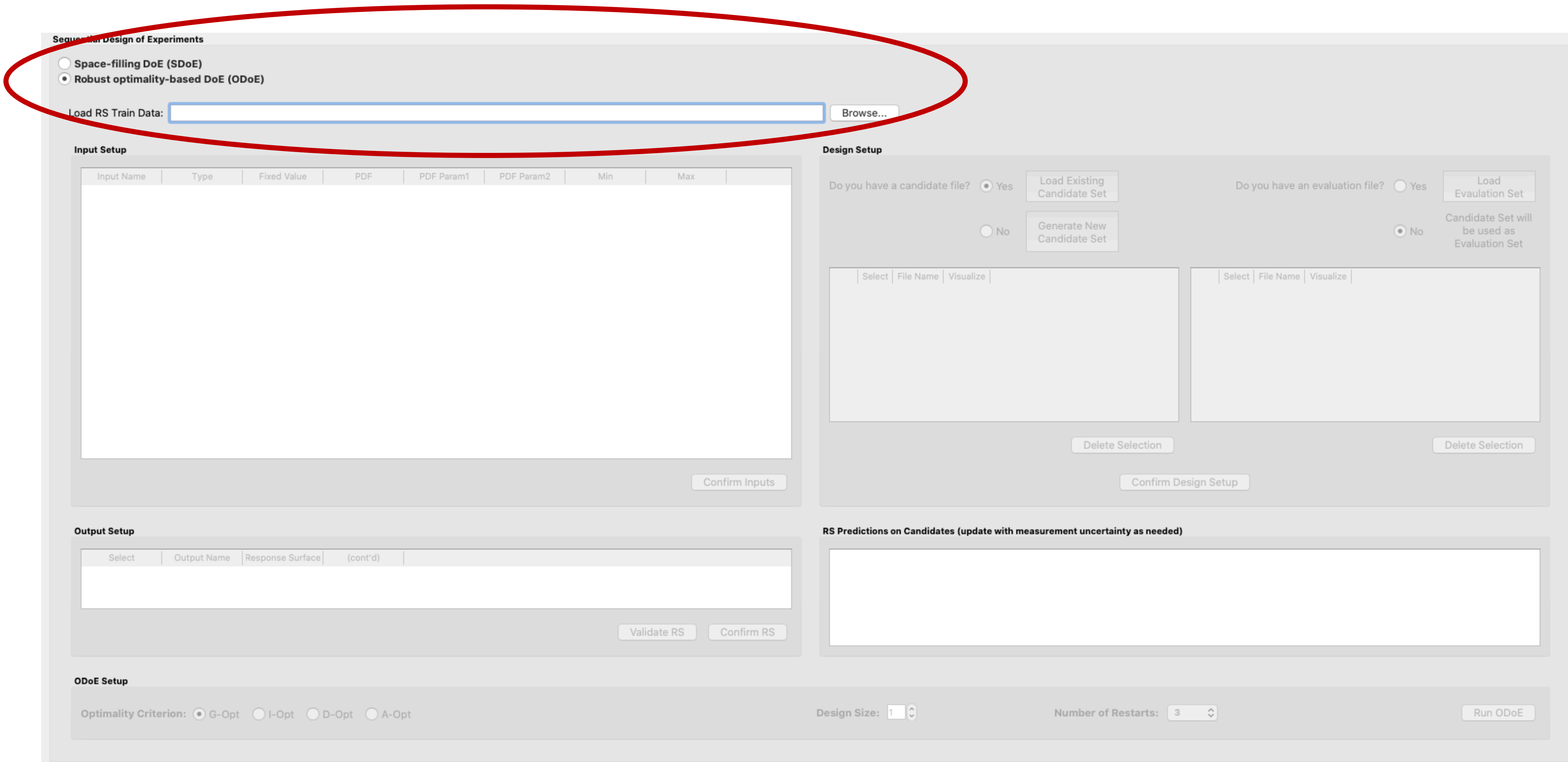


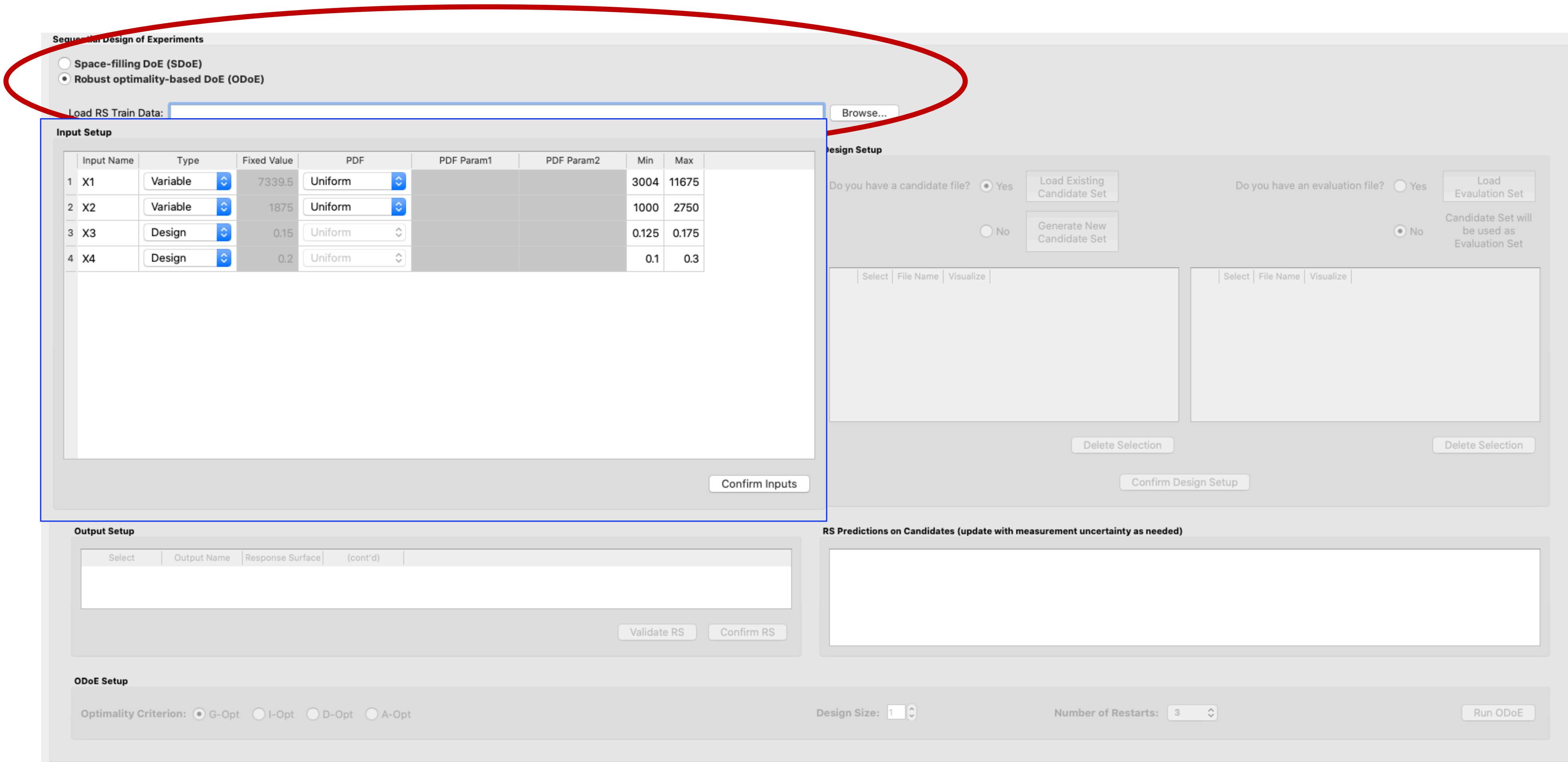
Designs
1
2
3
4
5
6
Input Distance: 0.3077
Response Distance: 0.5625
Pareto front
Maximin Response Distance
Maximin Input Distance
Frequency
Design Points
Frequency
Design Points
Frequency
Design Points
Frequency
Design Points
X1
X2
Y

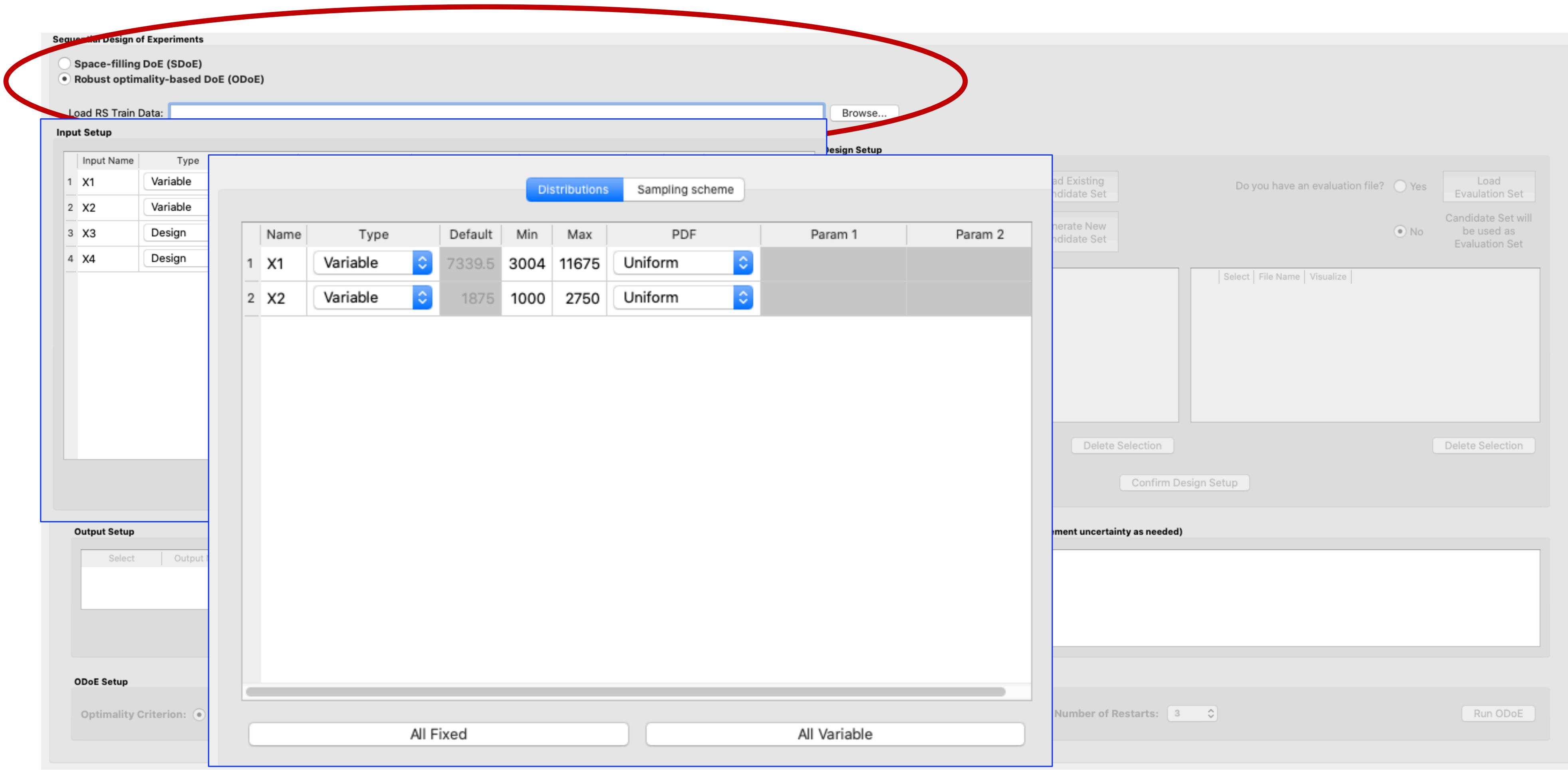
Robust Optimality-Based Design of Experiments (ODoE)

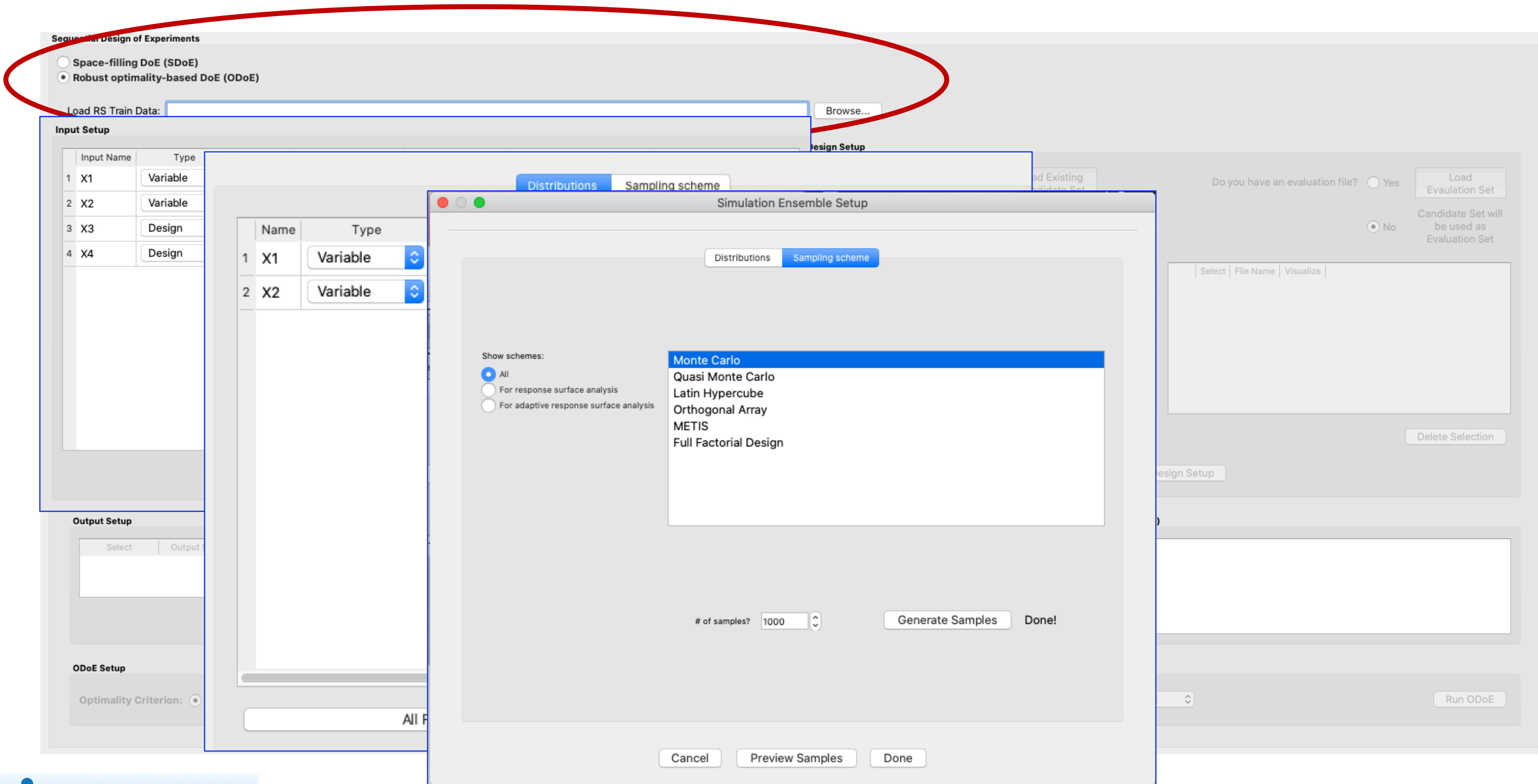
- **Goal:** construct ideal designs based on empirically fit models
- **Basic steps:**
  1. Select spreadsheet with columns of inputs and column(s) of responses
  2. Identify type of each input: Variable – not controllable during experiment
     Design – controllable during experiment
  3. Specify details for each input (ranges, distribution shape, etc)
  4. Specify candidate set and evaluation set (optional)
  5. Fit an empirical model (different forms available) between inputs and response
  6. Evaluate fit of model. When satisfied with fit, proceed.
  7. Use model to generate a design. Choices:
     - Optimality criterion: G-, I- (focus: prediction),
       D-, A- (focus: parameter estimation)
     - Design size
     - Number of Restarts


FOQUS -- [not saved yet]
Flowsheet
Uncertainty
Optimization
OUU
SDoE
Surrogates
Settings
Help
Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Design Setup
Generate New Candidate Set
Load Existing Set
Clone Selected
Delete Selected
Save Selected
File Type
Visualize
File Name
• If you have a candidate set:
1. Press Load Existing Set and upload. Set File Type to Candidate.
2. If you also have previous data, press Load Existing Set and upload. Set File Type to Previous Data.
• If you do not have a candidate set, press "Generate New Candidate Set"
• When finished identifying previous data and candidate sets, press Continue.
Continue
Design Construction
Descriptor
Visualize
Candidate File
Previous Experiments/Data
Output Directory
Design Method
Back to Selection
Open SDoE Dialog
Filtering
Select Variables (columns) and/or Sample Points (rows) for Deletion.
Type new values for outputs in the appropriate cells.
Reset Table
Perform Deletion then Save as New Ensemble
Working Directory: /Users/sotorrio1/PycharmProjects/FOQUS-sotorrio

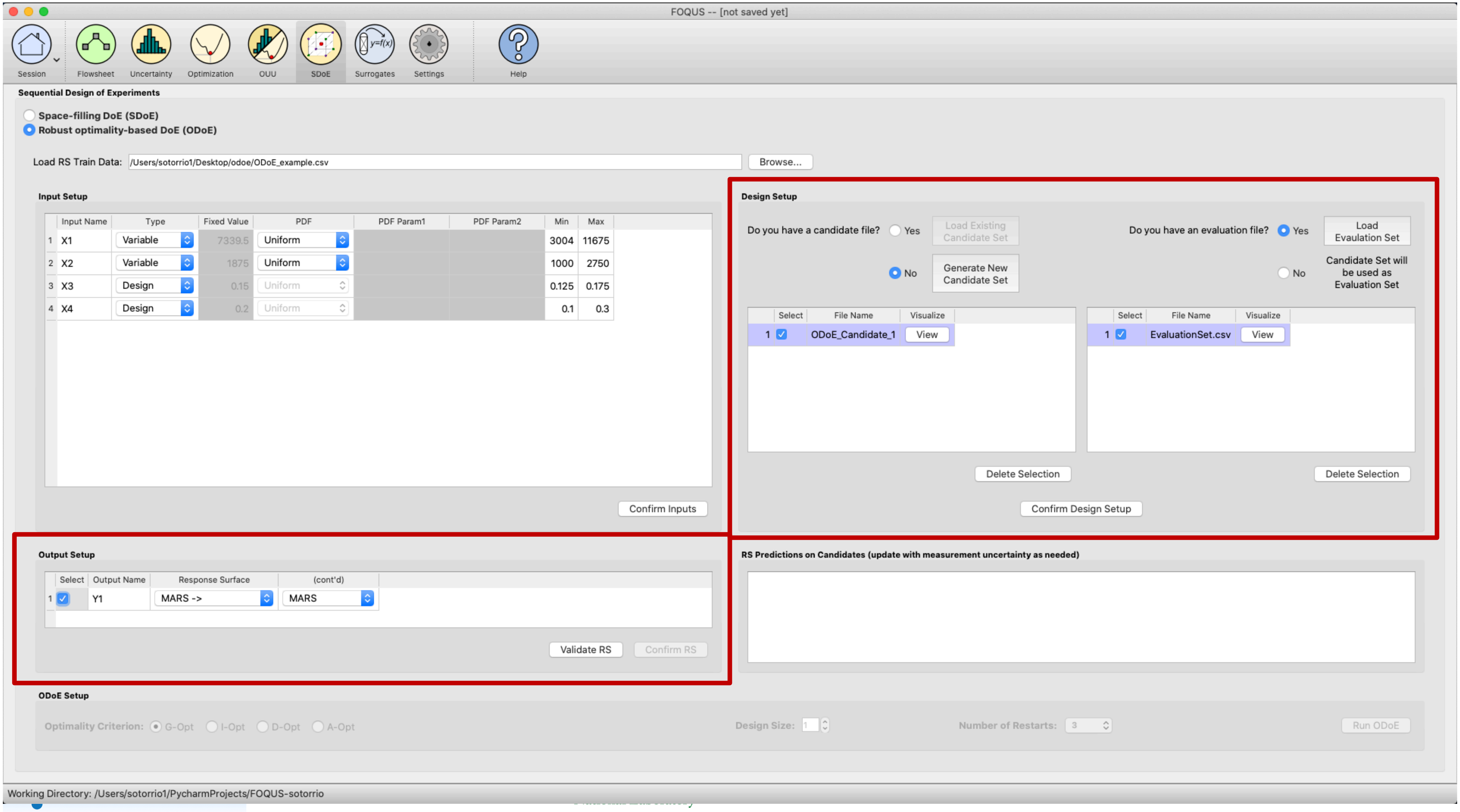

Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data:
Browse...
Input Setup
Input Name
Type
Fixed Value
PDF
PDF Param1
PDF Param2
Min
Max
Confirm Inputs
Design Setup
Do you have a candidate file?
Yes
No
Load Existing Candidate Set
Generate New Candidate Set
Do you have an evaluation file?
Load Evaulation Set
Candidate Set will be used as Evaluation Set
Select
File Name
Visualize
Delete Selection
Confirm Design Setup
Output Setup
Select
Output Name
Response Surface
(cont'd)
Validate RS
Confirm RS
RS Predictions on Candidates (update with measurement uncertainty as needed)
ODoE Setup
Optimality Criterion:
G-Opt
I-Opt
D-Opt
A-Opt
Design Size:
Number of Restarts:
Run ODoE


Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data:
Browse...
Input Setup
Input Name | Type | Fixed Value | PDF | PDF Param1 | PDF Param2 | Min | Max
1 X1 Variable 7339.5 Uniform 3004 11675
2 X2 Variable 1875 Uniform 1000 2750
3 X3 Design 0.15 Uniform 0.125 0.175
4 X4 Design 0.2 Uniform 0.1 0.3
Confirm Inputs
Design Setup
Do you have a candidate file? Yes No
Load Existing Candidate Set
Generate New Candidate Set
Do you have an evaluation file? Yes No
Load Evaluation Set
Candidate Set will be used as Evaluation Set
Select | File Name | Visualize
Delete Selection
Confirm Design Setup
Output Setup
Select | Output Name | Response Surface | (cont'd)
Validate RS
Confirm RS
RS Predictions on Candidates (update with measurement uncertainty as needed)
ODoE Setup
Optimality Criterion: G-Opt I-Opt D-Opt A-Opt
Design Size: 1
Number of Restarts: 3
Run ODoE

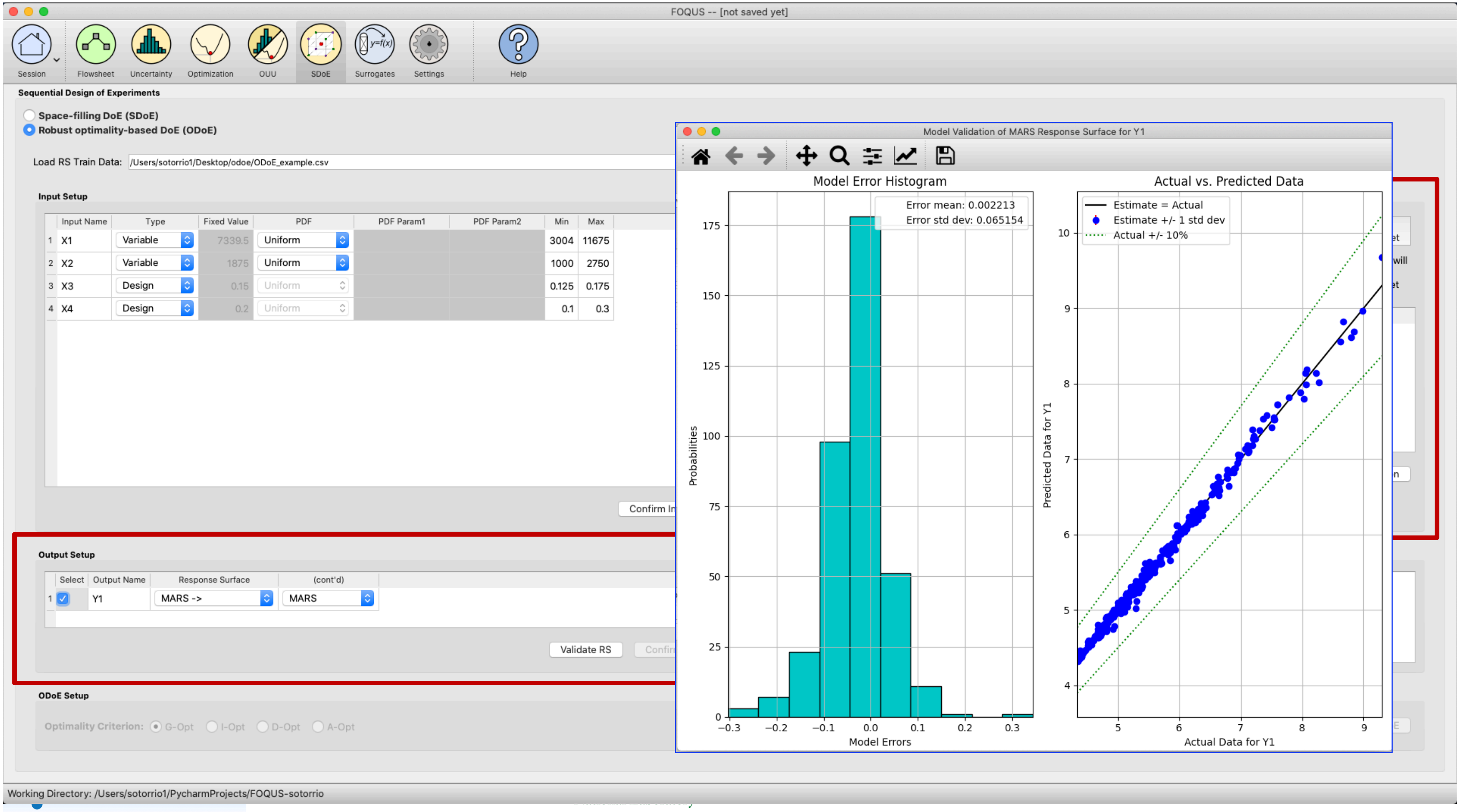

Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data:
Browse...
Input Setup
Input Name
Type
1 X1 Variable
2 X2 Variable
3 X3 Design
4 X4 Design
Distributions
Sampling scheme
Name | Type | Default | Min | Max | PDF | Param 1 | Param 2
1 X1 Variable 7339.5 3004 11675 Uniform
2 X2 Variable 1875 1000 2750 Uniform
All Fixed
All Variable
Output Setup
Select
ODoE Setup
Optimality Criterion:
Do you have an evaluation file?
Yes
No
Load Evaluation Set
Candidate Set will be used as Evaluation Set
Select
File Name
Visualize
Delete Selection
Delete Selection
Confirm Design Setup
Number of Restarts:
3
Run ODoE

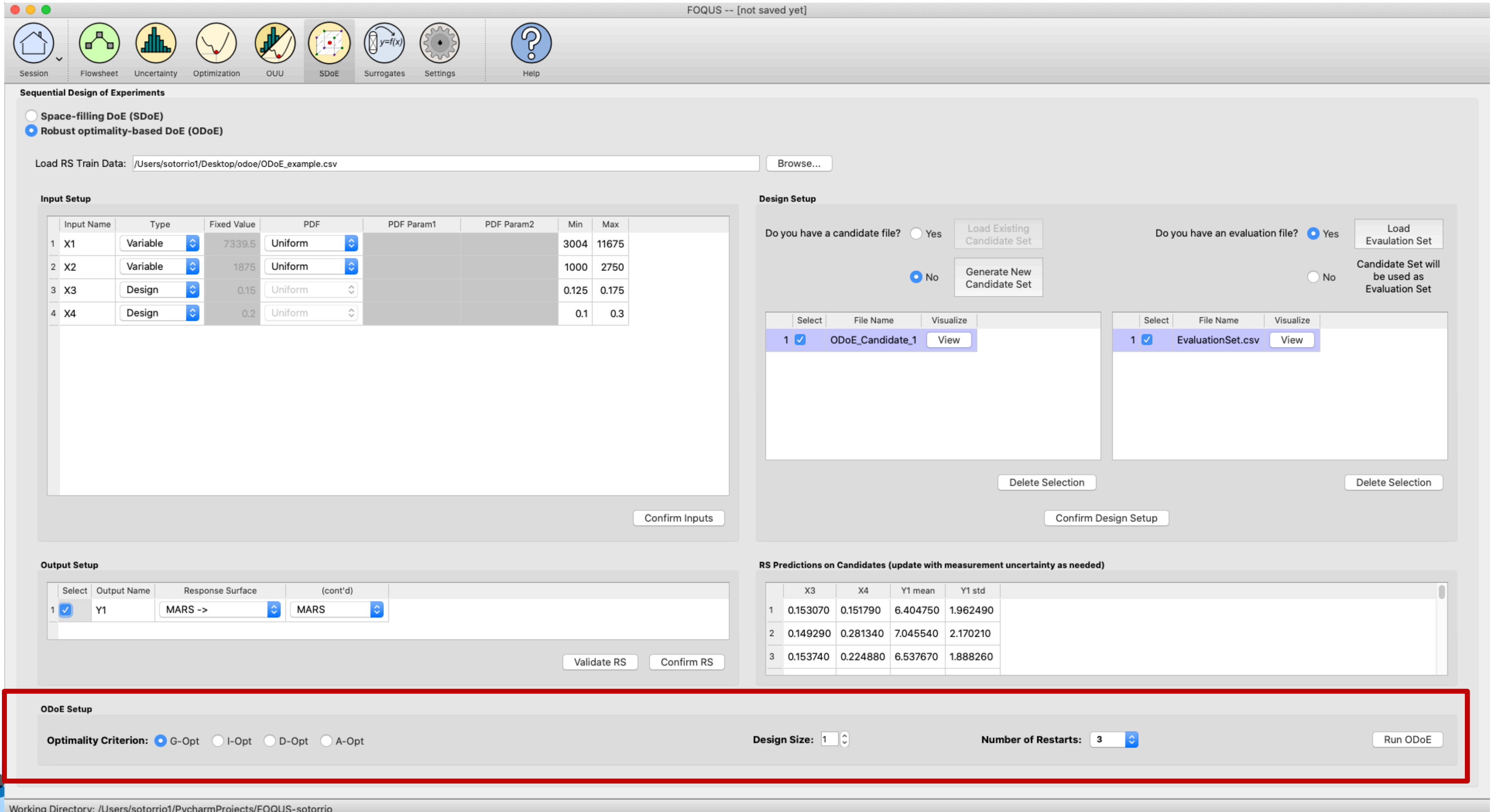


Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data:
Browse...
Input Setup
Input Name
Type
X1 Variable
X2 Variable
X3 Design
X4 Design
Output Setup
Select
ODoE Setup
Optimality Criterion:
Distributions
Sampling scheme
Name
Type
X1 Variable
X2 Variable
Simulation Ensemble Setup
Distributions
Sampling scheme
Show schemes:
All
For response surface analysis
For adaptive response surface analysis
Monte Carlo
Quasi Monte Carlo
Latin Hypercube
Orthogonal Array
METIS
Full Factorial Design
# of samples?
1000
Generate Samples
Done!
Cancel
Preview Samples
Done
Do you have an evaluation file?
Yes
No
Load Evaluation Set
Candidate Set will be used as Evaluation Set
Select
File Name
Visualize
Delete Selection
Run ODoE

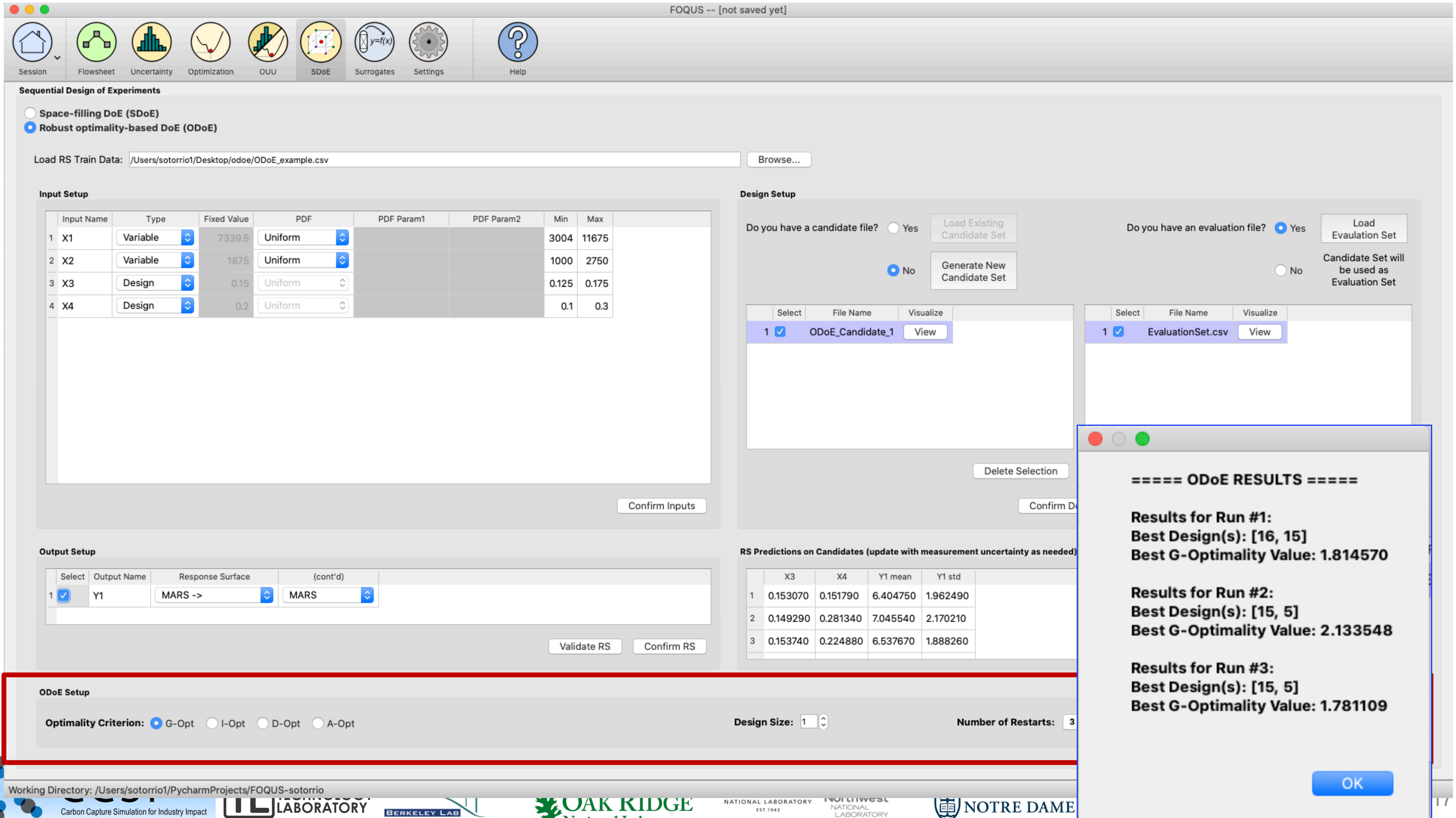


FOQUS -- [not saved yet]
Session
Flowsheet
Uncertainty
Optimization
OUU
SDoE
Surrogates
Settings
Help
Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data: /Users/sotorrio1/Desktop/odoe/ODoE_example.csv
Browse...
Input Setup
Input Name | Type | Fixed Value | PDF | PDF Param1 | PDF Param2 | Min | Max
1 X1 Variable 7339.5 Uniform 3004 11675
2 X2 Variable 1875 Uniform 1000 2750
3 X3 Design 0.15 Uniform 0.125 0.175
4 X4 Design 0.2 Uniform 0.1 0.3
Confirm Inputs
Design Setup
Do you have a candidate file?
Yes
Load Existing Candidate Set
No
Generate New Candidate Set
Do you have an evaluation file?
Yes
Load Evaluation Set
No
Candidate Set will be used as Evaluation Set
Select | File Name | Visualize
1 ODoE_Candidate_1 View
Select | File Name | Visualize
1 EvaluationSet.csv View
Delete Selection
Delete Selection
Confirm Design Setup
Output Setup
Select | Output Name | Response Surface | (cont'd)
1 Y1 MARS -> MARS
Validate RS
Confirm RS
RS Predictions on Candidates (update with measurement uncertainty as needed)
ODoE Setup
Optimality Criterion: G-Opt I-Opt D-Opt A-Opt
Design Size: 1
Number of Restarts: 3
Run ODoE
Working Directory: /Users/sotorrio1/PycharmProjects/FOQUS-sotorrio


FOQUS -- [not saved yet]
Session
Flowsheet
Uncertainty
Optimization
OUU
SDoE
Surrogates
Settings
Help
Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data: /Users/sotorrio1/Desktop/odoe/ODoE_example.csv
Input Setup
Input Name | Type | Fixed Value | PDF | PDF Param1 | PDF Param2 | Min | Max
1 X1 Variable 7339.5 Uniform 3004 11675
2 X2 Variable 1875 Uniform 1000 2750
3 X3 Design 0.15 Uniform 0.125 0.175
4 X4 Design 0.2 Uniform 0.1 0.3
Confirm In
Output Setup
Select | Output Name | Response Surface | (cont'd)
1 Y1 MARS -> MARS
Validate RS
Confirm
ODoE Setup
Optimality Criterion: G-Opt I-Opt D-Opt A-Opt
Working Directory: /Users/sotorrio1/PycharmProjects/FOQUS-sotorrio
Model Validation of MARS Response Surface for Y1
Model Error Histogram
Error mean: 0.002213
Error std dev: 0.065154
Probabilities
Model Errors
Actual vs. Predicted Data
Estimate = Actual
Estimate +/- 1 std dev
Actual +/- 10%
Predicted Data for Y1
Actual Data for Y1



Carbon Capture Simulation for Industry Impact


FOQUS -- [not saved yet]
Session
Flowsheet
Uncertainty
Optimization
OUU
SDoE
Surrogates
Settings
Help
Sequential Design of Experiments
Space-filling DoE (SDoE)
Robust optimality-based DoE (ODoE)
Load RS Train Data: /Users/sotorrio1/Desktop/odoe/ODoE_example.csv
Browse...
Input Setup
Input Name
Type
Fixed Value
PDF
PDF Param1
PDF Param2
Min
Max
1 X1 Variable 7339.5 Uniform 3004 11675
2 X2 Variable 1875 Uniform 1000 2750
3 X3 Design 0.15 Uniform 0.125 0.175
4 X4 Design 0.2 Uniform 0.1 0.3
Confirm Inputs
Design Setup
Do you have a candidate file?
Yes
No
Load Existing Candidate Set
Generate New Candidate Set
Do you have an evaluation file?
Yes
No
Load Evaluation Set
Candidate Set will be used as Evaluation Set
Select
File Name
Visualize
1 ODoE_Candidate_1 View
1 EvaluationSet.csv View
Delete Selection
Confirm D
Output Setup
Select
Output Name
Response Surface
(cont'd)
1 Y1 MARS -> MARS
Validate RS
Confirm RS
RS Predictions on Candidates (update with measurement uncertainty as needed)
X3
X4
Y1 mean
Y1 std
1 0.153070 0.151790 6.404750 1.962490
2 0.149290 0.281340 7.045540 2.170210
3 0.153740 0.224880 6.537670 1.888260
ODoE Setup
Optimality Criterion:
G-Opt
I-Opt
D-Opt
A-Opt
Design Size: 1
Number of Restarts: 3
Working Directory: /Users/sotorrio1/PycharmProjects/FOQUS-sotorrio
===== ODoE RESULTS =====
Results for Run #1:
Best Design(s): [16, 15]
Best G-Optimality Value: 1.814570
Results for Run #2:
Best Design(s): [15, 5]
Best G-Optimality Value: 2.133548
Results for Run #3:
Best Design(s): [15, 5]
Best G-Optimality Value: 1.781109
OK
Carbon Capture Simulation for Industry Impact
LABORATORY
OAK RIDGE
National Laboratory
NOTRE DAME

How this work fits into CCSI²

Computer Experiments
Physical Experiments – lab, pilot

Feasibility study → Exploration of Input Space → Model Building / Refinement → Optimization → Confirmation

Uniform SF

Non-Uniform SF (to reduce uncertainty)

Robust Optimal Design (to improve parameter estimation or prediction)

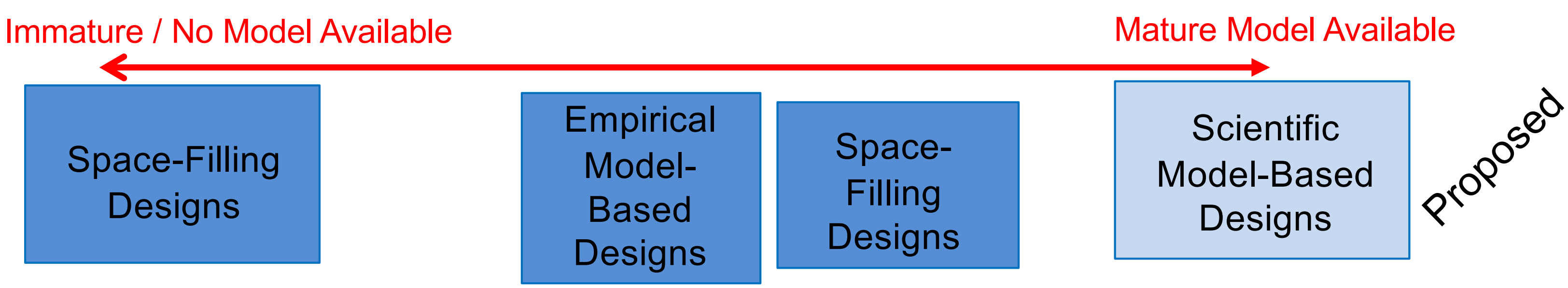
Non-Uniform SF (to emphasize near optimum)

Input-Response SF (to balance spread in input and response spaces)

CCSI²
Carbon Capture Simulation for Industry Impact

Future R&D Plan and Challenges

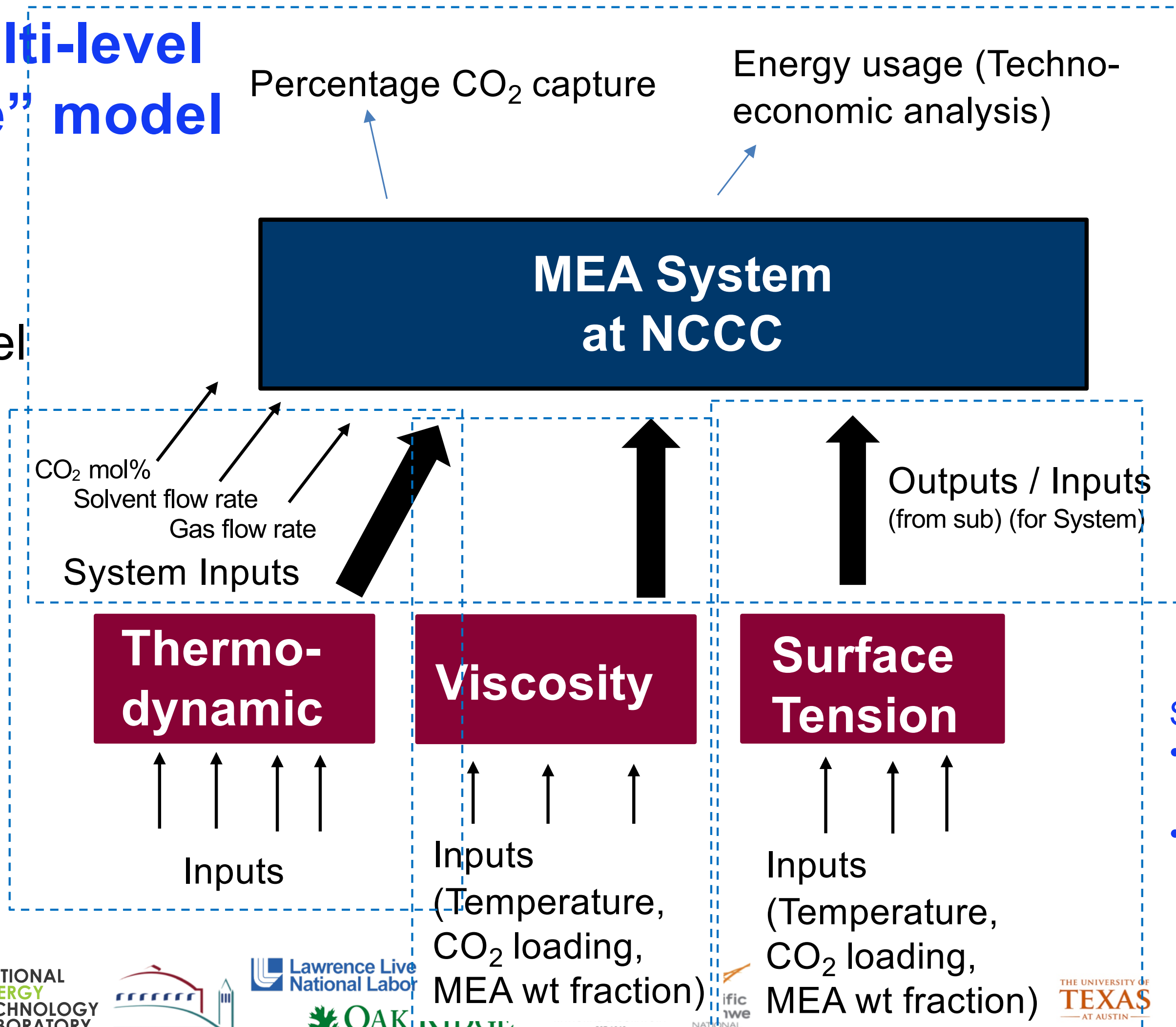
- There are still **big strategic opportunities** to pursue:
  1. Our current tools focus on the individual experiment level, but there are powerful opportunities when we **consider the big picture**
     - CCSI^2 supports development of systems and models spanning basic science to deployment. This involves science sub-system models, overall system model – each with their own associated costs and utility
     - When we consider multiple different experiments to achieve strategic goals for system performance, consider
       - what types of data should we be collecting?
       - how much of each type?
       - what design within each type?
  2. The CCSI^2 approach is **model-centric**. Additional design of experiments tools can leverage knowledge from a mature mathematical / science model



Design for multi-level "conglomerate" model

Engineering model of pilot system

Fundamental science models



Overall goal:
- Best prediction throughout the input space of NCCC system model

4 "experiments" to collect data, each with different
- Costs
- Inputs / outputs
- Utility

Scenarios to explore:
- Different levels of maturity of model
- Different available data

New Capability: Science Model-Based Design of Experiments

Main Idea: use full model equations <u>directly</u> to optimizing experimental campaigns to improve parameter estimates

+ Avoids need to build/validate surrogates
+ Discriminate between alternative mechanistic models
- Requires access to equations (e.g., Pyomo)

EY 2021 Progress (Sub-Task 2.1 & 3.1)

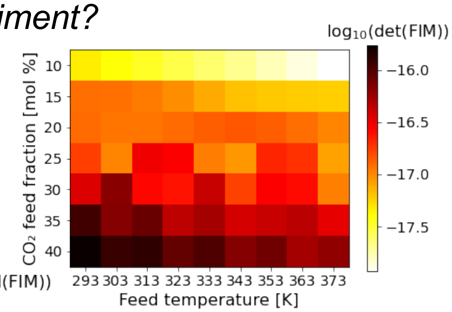
- Created MBDOE framework that works with any Pyomo model
- Demonstrated capability in DoE case study for fixed-bed MOF characterization

EY 2022 Proposed Work (Sub-Task 3.1)

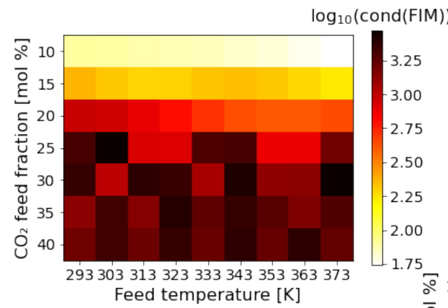
- Release framework open source as Pyomo package
- Create plan to integrate MBDOE in FOQUS
- Algorithm improvements to increase speed & robustness

Example: *What is the optimal CO_2 feed composition and feed temperature for next MOF fixed bed experiment?*

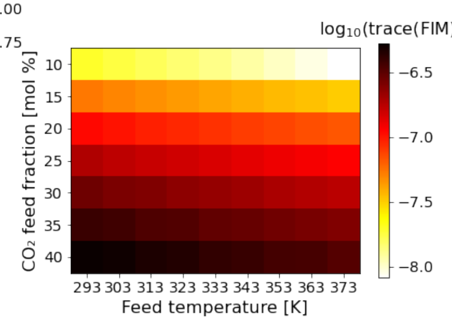
D-optimality →



← A-optimality



E-optimality →



CCSI2 Carbon Capture Simulation for Industry Impact

Planned work for EY21-EY23

- EY21
  - Missing values / Imputation for NUSF and IRSF (LLNL)
  - Videos and Documentation update for all capabilities (LANL & LLNL)
  - Robust optimality-based DoE enhancement (LLNL)
  - Science-based optimal design methodology development (ND)
  - Design for Conglomerate model methodology development and demonstration (LANL)
  - Collaboration with Pilot project teams – write up case study (LANL)
- EY22
  - Integration of Science-based optimal design into toolset
  - Integration of Design for Conglomerate model into toolset
  - Collaboration with Pilot project teams – write up case study
  - Update supporting materials (video, documentation)
- EY23
  - Collaboration with Pilot project teams – write up case study

Proposed for Breakout Discussion

- **Breakout Discussion**
  - Design for Multi-level "conglomerate" models
  - Design capability for Science-based models
  - Description of planned supporting materials – documentation, videos
  - Where do we anticipate design of experiments support being needed in the coming years?
    - Pilot studies: RTI, MTR, TDA
    - DOCCSS: PNNL CO2BOL, LBNL MOF
    - Other?
  - Wishlist for other design capabilities

- **Open Q&A**
  - ??

CCSI² Carbon Capture Simulation for Industry Impact

Acknowledgements

Co-Lead: Brenda Ng, LLNL

<u>Team:</u>
Pedro Sotorrio (LLNL),
Abby Nachtsheim (LANL),
Miranda Martin (LANL),
Alex Dowling (Notre Dame),
Jialu Wong (Notre Dame),
Josh Morgan (NETL),
Charles Tong (LLNL)

CCSI²

Carbon Capture Simulation for Industry Impact

For more information
https://www.acceleratecarboncapture.org/

Contact Info:

Christine Anderson-Cook, LANL
candcook@lanl.gov
505-695-8850

Brenda Ng, LLNL
ng30@llnl.gov